

Sistema de Recomendação de Pontos de Interesse em Recife com Filtragem Baseada em Conteúdo e Passaporte Pernambuco

Douglas I. F. Ferreira¹, Everton Gomedes²

¹Escola Superior de Agricultura Luiz de Queiroz (Esalq), Pecege
Universidade de São Paulo (USP) – Piracicaba – SP – Brasil

²Faculdade de Engenharia Elétrica e de Computação
Universidade Estadual de Campinas (FEEC/UNICAMP) – Campinas – SP – Brasil

douglas.iff@usp.br, gomedes@unicamp.br

Abstract. *Recife has significant tourism potential but requires digital solutions to promote its Points of Interest (POIs). Thus, implementing an intelligent system represents an innovation opportunity for the sector. A Content-Based Recommendation System was proposed, which uses similarity metrics to correlate user preferences with POI characteristics. The system achieved 79% precision, 100% minimum hit rate, and 86% coverage, demonstrating strong performance in cold start scenarios despite the overexposure of some POIs. Therefore, the model proved effective in providing personalized recommendations, showing potential to boost local tourism.*

Resumo. *Recife tem grande potencial turístico, mas necessita de soluções digitais para promover seus Pontos de Interesse (POIs). A implementação de um sistema inteligente representa uma oportunidade de inovação no setor. Foi proposto um Sistema de Recomendação via Filtragem Baseada em Conteúdo, que usa TF-IDF e métricas de similaridade para correlacionar as preferências dos usuários com as características dos POIs. O sistema obteve 79% de precisão, 100% de taxa mínima de acerto e 86% de cobertura, demonstrando bom desempenho em cenários de cold start, apesar da superexposição de alguns POIs. Sendo assim, o modelo mostrou eficácia na recomendação personalizada, apresentando potencial para impulsionar o turismo local.*

1. Introdução

Um grande volume de informações é gerado diariamente, levantando um desafio crítico: filtrar conteúdos relevantes para o usuário [Li et al. 2021]. Nesse contexto, os Sistemas de Recomendação (SRs) surgem como uma solução eficaz para sugerir itens alinhados às preferências dos usuários, contribuindo para a mitigação da sobrecarga informacional. Plataformas como *Netflix*, *Amazon* e *Tripadvisor* exemplificam a aplicação bem-sucedida de sistemas altamente personalizados [Aggarwal 2016].

No turismo, os SRs são utilizados para oferecer sugestões sobre transporte, hospedagem e atrações, Pontos de Interesse (POIs), redes sociais, rotas e planos de viagem [Borràs et al. 2014, Sarkar et al. 2023]. Apesar dos avanços, ainda há desafios tecnológicos em grandes centros urbanos brasileiros, como Recife, que demandam soluções integradas para conectar os SRs aos seus respectivos POIs.

Análises em bases de dados como Google Acadêmico, SciELO e Periódicos da CAPES revelam uma lacuna na literatura sobre o tema, especialmente no contexto brasileiro. A implementação de SRs de POIs representam uma oportunidade de inovação digital no setor turístico, especialmente no cenário pós-pandemia da Covid-19 [Ferreira and Ramos 2024, Trigo 2020].

Recife, capital de Pernambuco e objeto de estudo, tem investido em iniciativas que promovem a adoção de soluções tecnológicas no turismo, destacando o potencial dessa área [Campos et al. 2022]. A cidade possui importantes POIs, muitos com valor histórico e acesso gratuito, além de uma diversidade de festividades culturais. Essa variedade, embora vantajosa, intensifica a sobrecarga informacional enfrentada pelos turistas. Alguns atrativos estão integrados no Passaporte Pernambuco, um documento não oficial que estimula o turismo no estado ao permitir o registro simbólico das visitas [Ferreira and Ramos 2024, Fonseca 2020].

Diante da lacuna na literatura e da relevância do turismo regional, este estudo visa investigar a implementação de técnicas de recomendação de POIs na política pública do Passaporte Pernambuco, utilizando a Filtragem Baseada em Conteúdo (FBC). Essa abordagem permite personalizar sugestões com base em atributos descritivos dos POIs e nas preferências individuais dos usuários [Aggarwal 2016].

Este estudo baseia-se nas seguintes hipóteses: (i) um Sistema de Recomendação (SR) utilizando FBC pode gerar recomendações relevantes, mesmo em contextos de dados limitados (*cold start*); e (ii) a personalização das recomendações de POIs conforme as preferências do usuário pode aumentar a cobertura e a diversidade dos itens sugeridos.

Este trabalho está estruturado da seguinte forma: a Seção 2 aborda os fundamentos dos SRs, com foco nos aplicados a POIs; a Seção 3 detalha o modelo de FBC; a Seção 4 apresenta o contexto e os objetivos da política pública do Passaporte Pernambuco. Em seguida, são descritos os materiais e métodos utilizados, seguidos pela apresentação dos resultados e, por último, pelas considerações finais.

2. Trabalhos Relacionados

SRs são softwares que utilizam métodos estatísticos e técnicas de aprendizado de máquina. O principal objetivo dos SRs é sugerir itens alinhados às preferências individuais dos usuários, com base nas informações de seus perfis. Esses perfis contêm dados relevantes que permitem ao sistema oferecer recomendações mais precisas e personalizadas [Aggarwal 2016, Castells and Jannach 2023, Gao et al. 2010].

Os SRs podem ser classificados em dois modelos principais. O modelo de (i) previsão visa estimar a compatibilidade entre um usuário e um item, utilizando algoritmos de aprendizado treinados com dados de históricos de interações. Em contrapartida, o modelo de (ii) classificação foca na seleção direta dos itens mais relevantes, com base nas características do perfil do usuário [Aggarwal 2016].

Além disso, os SRs utilizam (i) dados de interação, que registram comportamentos como avaliações, cliques e compras; e (ii) dados de atributos, que descrevem usuários e itens por meio de perfis, descrições textuais e palavras-chave. Esses dados são processados por diferentes abordagens de recomendação, destacando-se o modelo colaborativo, o modelo baseado em conteúdo e os modelos híbridos [Aggarwal 2016, Borràs et al. 2014,

Castells and Jannach 2023, Li et al. 2021].

Para avaliar o desempenho das recomendações, são coletadas avaliações contínuas, intervalares, ordinais, binárias ou unárias. Podendo ser expressas de forma explícita (interação direta) ou implícita (quando derivadas de comportamentos) [Aggarwal 2016].

Recentemente, destacam-se o uso de *Large Language Models* (LLMs), como *DeepSeek*, para interagir com aplicações. O uso das LLMs demonstra um grande potencial de revolucionar os SRs por compreenderem melhor a linguagem e intenções humanas com maior naturalidade [Zhao et al. 2024]. Além de SRs baseados em *Deep Learning*, como os *Autoencoders*, que têm demonstrado alta eficácia na personalização de recomendações para domínios complexos [Gomede et al. 2021]

No turismo, os SRs são essenciais para personalizar sugestões de destinos, atividades e serviços. A aplicação das abordagens de filtragem e avaliação discutidas pode ser adaptada, um exemplo são os SRs de POIs, que se destacam como ferramentas relevantes na conexão entre turistas e atrações locais.

2.1. Sistemas de Recomendação de Pontos de Interesse

Os SRs de POIs visam sugerir novos locais para exploração pelos usuários, especialmente diante do aumento contínuo de POIs disponíveis. Além de facilitar a descoberta de destinos, essas recomendações elevam a probabilidade de uma experiência turística satisfatória, promovendo a integração entre os ambientes físico e digital [Zhao et al. 2016].

A recomendação de POIs pode ser dividida em dois objetivos principais: (i) a Recomendação Geral, que sugere os *top-k* POIs com base nas preferências globais do usuário; e (ii) a Recomendação Sucessiva, que considera o histórico de locais visitados para identificar e recomendar POIs com características semelhantes, explorando relações de similaridade [Zhao et al. 2016].

Neste contexto, a informação baseada em conteúdo é fundamental, pois permite estabelecer relações entre as preferências dos usuários e as características dos POIs [Gao et al. 2015]. Segundo a literatura, os SRs de POIs fundamentados na abordagem de conteúdo dependem de dois conjuntos de atributos: os dos POIs, que incluem categoria, localização e descrição textual, e os dos usuários, que abrangem interesses e comportamentos anteriores [Kang et al. 2006].

O objetivo deste projeto é selecionar itens relevantes para o perfil de cada usuário, utilizando a técnica de FBC. A próxima seção apresenta uma análise detalhada dessa abordagem, destacando seus fundamentos e aplicações no contexto do turismo.

3. Filtragem Baseada em Conteúdo

A FBC utiliza informações descritivas para recomendar itens ao usuário. Essa abordagem combina as preferências, expressas ou inferidas, do usuário com as características dos itens disponíveis, gerando recomendações personalizadas [Aggarwal 2016, Borràs et al. 2014, Castells and Jannach 2023, Li et al. 2021].

Este modelo oferece vantagens significativas em contextos de escassez de dados, como nos problemas de *cold-start* de itens e de usuários. Nesses casos, os métodos de filtragem colaborativa costumam ter desempenho inferior devido à falta de informações.

Por outro lado, a FBC apresenta limitações, especialmente em relação à diversidade: como as recomendações se baseiam apenas em características conhecidas, o sistema tende a sugerir itens muito semelhantes aos já consumidos, resultando em recomendações previsíveis e repetitivas [Aggarwal 2016, Castells and Jannach 2023, Li et al. 2021].

O ranqueamento dos itens candidatos é baseado na distância ou similaridade entre as descrições dos itens e o perfil do usuário, que pode ser construído a partir do histórico de interações ou de um modelo explícito de preferências [Li et al. 2021]. Para representar as informações descritivas, utiliza-se técnicas de vetorização textual, como o *Term Frequency–Inverse Document Frequency* (TF-IDF). O algoritmo do TF-IDF é simples e eficiente, sendo uma técnica numérica de estatística que gera o resultado da relevância de um termo em um conjunto de documentos [Salton and Buckley 1988, Qaiser and Ali 2018].

A comparação entre os vetores gerados pelo TF-IDF é realizada por meio de métricas de similaridade, incluindo Distância Euclidiana (ED), Similaridade do Cosseno (COS), Coeficiente de Correlação de Pearson (PCC) e Coeficiente de Jaccard (JC) [Castells and Jannach 2023, Magara et al. 2018, Salton and Buckley 1988, Sondur and Chigadani 2016]. Os algoritmos avaliam as similaridades entre as preferências do usuário e a descrição dos POIs, sendo uma abordagem eficaz para personalizar recomendações turísticas, especialmente no contexto do Passaporte Pernambuco, que será apresentado na próxima seção.

4. Passaporte Pernambuco

Recife possui uma economia diversificada, destacando-se recentemente o setor de turismo, que tem se consolidado como um eixo estratégico para o desenvolvimento local [Moraes et al. 2022]. Apenas no carnaval de 2024, a cidade recebeu mais de 3,4 milhões de visitantes, impactando a economia em cerca de 2,4 bilhões de reais [Costa 2024].

Nos últimos anos, Recife tem adotado práticas de turismo criativo e sustentável, assim como demonstrou uma demanda por inovação digital, promovendo uma "nova forma de fazer turismo". Essa abordagem valoriza a imersão cultural dos visitantes, incentivando experiências autênticas e interativas que fortalecem os laços entre os turistas e a identidade local [Campos et al. 2022, Ferreira and Ramos 2024].

Dessa forma, os POIs deste SR foram extraídos do Passaporte Pernambuco, lançado em 2020 para promover o turismo local. Com esse passaporte, os visitantes podem registrar simbolicamente suas visitas aos atrativos turísticos das cidades pernambucanas, carimbando as páginas correspondentes a cada POI visitado. O passaporte oferece atrativos como restaurantes, igrejas, museus, teatros e embaixadas [Fonseca 2020, Valença 2022].

5. Materiais e Métodos

Este estudo utilizou uma abordagem mista. A abordagem qualitativa ajudou a definir o domínio de aplicação do SR, por meio de uma análise da literatura e da compreensão do contexto do Passaporte Pernambuco. A abordagem quantitativa, por sua vez, tratou dados numéricos, incluindo o cálculo de similaridades e a aplicação de métricas para avaliar o sistema [Creswell and Creswell 2021].

A pesquisa é considerada de natureza aplicada, visando gerar um impacto prático na resolução de um problema específico. O foco está na prototipagem e validação técnica do

sistema por meio de um Mínimo Produto Viável (MVP) [Ries 2012], cujo código, equações, banco de dados, documentação e *prompts* estão disponíveis no *GitHub*¹ para reprodução. No campo da computação, a investigação é exploratória, destacando a importância da aplicação de técnicas no sistema e em seu processo de desenvolvimento [Wazlawick 2009].

Para calcular as métricas de similaridade da FBC, utilizou-se uma combinação de TF-IDF, COS, ED, JC e PCC [Gao et al. 2010, Qaiser and Ali 2018, Salton and Buckley 1988, Sondur and Chigadani 2016]. Por fim, foram utilizadas as métricas de avaliação: Precisão@k, *Hit Rate*@k, Cobertura, Similaridade Intra-Lista e Frequência de POI [Castells and Jannach 2023, Deshpande and Karypis 2004, Gomedes et al. 2021, Jesse et al. 2023, Li et al. 2021].

Para a geração de dados, foram utilizados os 36 POIs do Passaporte Pernambuco em Recife e criados 30 usuários fictícios utilizando LLMs. Esta abordagem pode ser caracterizada como uma simulação social simples baseada em *prompts*, essa técnica é considerada promissora para estudos exploratórios e validação de protótipos. É importante notar que este método possui limitações inerentes, como falta de diversidade e tender a criar perfis estereotipados [Anthis et al. 2025]. Entretanto, no contexto deste trabalho, consideramos esta abordagem adequada. Dessa forma, essas métricas permitiram uma análise dos resultados da implementação do sistema, que será apresentada na seção seguinte.

6. Implementação do Sistema

O foco principal do MVP foi a Recomendação Geral de POIs, utilizando um modelo de classificação baseado em conteúdo, fundamentado nos atributos descritivos dos POIs e no perfil do usuário. Como ilustrado na Figura 1, o fluxo do sistema começa com o cadastro do usuário e seus interesses, além do registro dos POIs no banco de dados. O usuário fornece informações gerais, como nome e idade, e interesses específicos, como motivações para viagem, hobbies e temas de interesse. Esses dados são armazenados com atributos estáticos (dados pessoais) e dinâmicos (preferências e histórico de interações).

Os POIs, como museus, igrejas e restaurantes, são registrados com atributos como nome, descrição e categorias. Quando o usuário solicita uma recomendação, o sistema compara seu perfil com as descrições dos POIs utilizando a técnica TF-IDF e métricas de similaridade, retornando os 5 POIs mais relevantes com base na similaridade dos termos.

Após receber as recomendações, o usuário pode avaliá-las de forma binária, atribuindo um *like* ou *dislike*. Essas avaliações são registradas no banco de dados e utilizadas no cálculo de métricas de avaliação do sistema, como Precisão@k e *Hit Rate*@k. Outras métricas, como Cobertura, Similaridade Intra-Lista e Frequência de POI, são calculadas com base nos dados gerados pelas recomendações.

Na Tabela 1, apresenta-se um exemplo de resultado do SR. A FBC nesse trabalho se diferencia por utilizar uma lista (*strings*) de termos estruturados de motivações, hobbies e temas, no qual o TF-IDF transforma em vetores numéricos. A dimensão do TF-IDF é fixada em 45, sempre utilizando todos os termos (15 de cada categoria). Vale destacar que no cálculo do IDF é utilizado a Suavização de Laplace, evitando o problema de probabilidade zero; e aplica-se a normalização vetorial antes de calcular a similaridade, conforme sugerido pela literatura [Salton and Buckley 1988].

¹<https://github.com/douglasfragoso/recommender-pe>

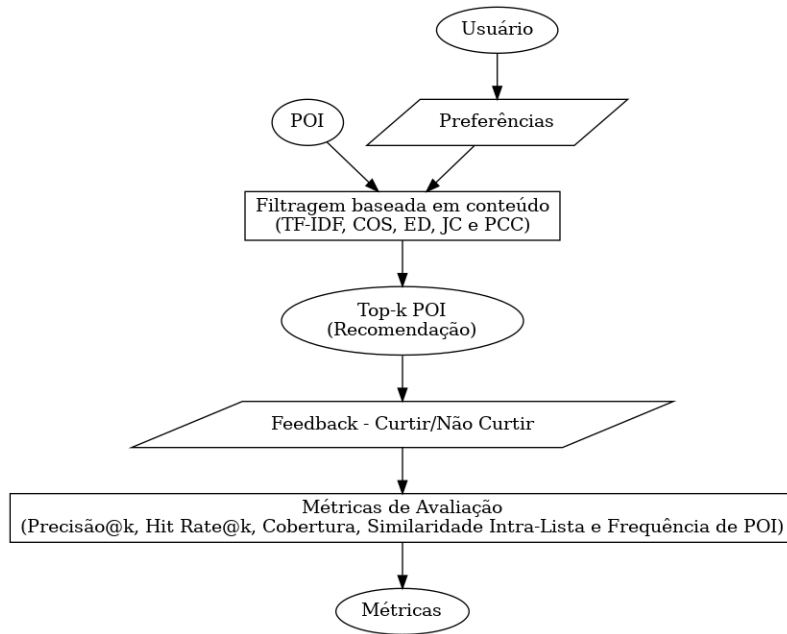


Figura 1. Fluxo do Sistema

Tabela 1. Exemplo de recomendação por métrica de similaridade

POI ID	Cosseno	Euclidiana	Pearson	Jaccard	Média
4	0.6019	0.5284	0.7534	0.3416	0.5563
36	0.5588	0.5156	0.7201	0.3032	0.5244
7	0.3601	0.4692	0.5891	0.1749	0.3983
29	0.3681	0.4708	0.6068	0.1454	0.3978
30	0.3681	0.4708	0.6068	0.1454	0.3978

A comparação entre o vetor do usuário e os vetores dos POIs é realizada por um híbrido de métricas de similaridade. A pontuação final de recomendação é a média aritmética das quatro métricas normalizadas para o intervalo (0, 1). Essa abordagem combinada captura diferentes facetas da similaridade entre o perfil do usuário e cada item, resultando em um ranqueamento mais holístico. Isto é, diminuiu a influência de vieses específicos de cada métrica [Magara et al. 2018, Sondur and Chigadani 2016].

As métricas PCC e COS geralmente obtiveram os valores mais altos, pois ignoram os pesos dos termos nos vetores, focando na direção dos vetores. ED apresentou valores intermediários, considerando essas diferenças de intensidade. Enquanto o JC teve os valores mais baixos, sendo o único a penalizar a ausência de termos em comum. Esses resultados convergem com os trabalhos relacionados ao uso de múltiplas métricas [Magara et al. 2018, Sondur and Chigadani 2016]. Com o MVP pronto, foram utilizados dados simulados por LLMs para avaliar o desenvolvimento do sistema. Essa etapa da pesquisa será descrita na próxima seção.

7. Coleta de Dados e Métricas de Avaliação

Foram utilizados os 36 POIs do Passaporte Pernambuco em Recife, junto com 30 usuários fictícios (com suas preferências e avaliações), criados por LLMs (*DeepSeek*). Esses dados

Tabela 2. Frequência de recomendação por POI

POI ID	Freq.	POI ID	Freq.	POI ID	Freq.	POI ID	Freq.
1	0.10	10	0.13	19	0.10	28	0.37
2	0.10	11	0.03	20	0.27	29	0.13
3	0.10	12	0.10	21	0.23	30	0.13
4	0.20	13	0.00	22	0.23	31	0.10
5	0.03	14	0.10	23	0.00	32	0.33
6	0.20	15	0.00	24	0.00	33	0.00
7	0.13	16	0.07	25	0.03	34	0.20
8	0.27	17	0.20	26	0.17	35	0.10
9	0.03	18	0.37	27	0.17	36	0.27

visam simular comportamentos de turistas e mostraram-se adequados para a validação técnica inicial do algoritmo proposto. Por natureza, as LLMs são treinadas com textos humanos, o que permite a capacidade de imitar comportamentos e opiniões humanas em diversos cenários [Anthis et al. 2025] .

Cada usuário recebeu uma lista personalizada de POIs recomendados com base em seu perfil e interações, registradas no banco de dados por meio de avaliações binária do tipo *like/dislike*. Esses registros foram utilizados para calcular as métricas de avaliação discutidas anteriormente.

Com os dados simulados, foram extraídas estatísticas básicas e identificadas tendências no comportamento do algoritmo. O SR apresentou uma precisão média de 79% na relevância dos POIs recomendados, com um Intervalo de Confiança (IC) de 95% entre 73% e 85%. Além disso, 100% dos usuários receberam pelo menos um POI relevante.

Em termos de cobertura de itens, 86% dos POIs foram recomendados ao menos uma vez, e a similaridade média intra-lista entre os itens recomendados foi de 41%. Embora os dados utilizados sejam simulados, os resultados foram considerados sólidos: o SR apresentou um bom equilíbrio entre engajamento e diversidade, evitando redundância excessiva nos itens sugeridos e mantendo alta precisão nas recomendações.

Como mostrado na Tabela 2, o sistema deixou de recomendar 13,88% dos POIs, com os itens 13, 15, 23, 24 e 33 apresentando 0% de ocorrência nas listas recomendadas. Em contrapartida, alguns POIs se destacaram pela popularidade excessiva: os itens 8, 18, 20, 28, 32 e 36 foram recomendados entre 27% e 37% das vezes, alcançando uma popularidade de 16,66%.

A superexposição e subutilização de certos POIs podem resultar na saturação de recomendações, prejudicando a experiência personalizada de turistas à medida que suas preferências mudam. Essas distorções estão possivelmente ligadas à cobertura desigual de categorias e ao viés algorítmico decorrente da similaridade textual.

Os dados simulados com LLMs permitiram observar o comportamento típico de um SR baseado em conteúdo [Aggarwal 2016, Castells and Jannach 2023, Li et al. 2021], evidenciando seus pontos fortes e limitações: desempenho satisfatório em contextos de *cold start*, mas tendência à repetição de certos itens nas recomendações. Esse cenário ressaltou a necessidade de adotar estratégias que ampliem a cobertura de itens e promovam

maior equilíbrio nas recomendações.

Como alternativa, propõe-se a introdução de pesos ajustáveis por termo de interesse, o que permitirá maior controle sobre o balanceamento das recomendações. Essa estratégia pode reduzir a predominância de POIs populares nos *rankings*, promovendo uma distribuição mais equitativa e personalizada dos itens sugeridos. Além disso, destacam-se as necessidades de expansão do conjunto de POIs e termos associados, assim como a incorporação de novas *features* (filtragem híbrida e avaliações textuais).

Por fim, a adoção da técnica de *Latent Semantic Analysis* (LSA), por meio da *Singular Value Decomposition* (SVD), é uma alternativa promissora para identificar relações semânticas mais profundas [Achakulvisut et al. 2016] entre as preferências dos usuários e as descrições dos POIs. Essa abordagem pode ampliar a descoberta de itens menos evidentes, que, embora não compartilhe termos explícitos, possuem similaridade temática.

8. Conclusão e Trabalhos Futuros

Este trabalho teve como objetivo investigar a implementação de técnicas de recomendação de POIs na política pública do Passaporte Pernambuco. Para isso, foi realizado um estudo abrangente sobre SRs, SRs aplicados a POIs e a técnica de FBC, além da análise da política pública relacionada ao Passaporte Pernambuco.

A pesquisa utilizou uma metodologia mista, com abordagem exploratória e natureza aplicada, resultando na implementação do MVP. O algoritmo desenvolvido empregou técnicas estatísticas de similaridade vetorial, como TF-IDF, COS, ED, JC e PCC. Para avaliar a qualidade das recomendações, foram utilizadas as métricas Precisão@k, *Hit Rate*@k, Cobertura, Similaridade Intra-Lista e Frequência de POI.

O sistema utilizou 36 POIs localizados em Recife, presentes no Passaporte Pernambuco, e dados simulados de 30 usuários gerados por meio de LLMs. A precisão média alcançada foi de 79%, com um IC de 95%, variando entre 73% e 85%, além de um *Hit Rate* de 100%. Observou-se que 86% dos POIs foram recomendados ao menos uma vez e a similaridade intra-lista entre os itens recomendados no *top-5* foi de 41%.

Os resultados indicam que o sistema teve um desempenho compatível com o previsto na literatura, especialmente em cenários de *cold start*. No entanto, notou-se uma certa repetição nas recomendações, atribuída à limitação do número de POIs e à semelhança de seus atributos descritivos.

As limitações da versão atual do MVP, embora esperadas, indicam caminhos claros para melhorias futuras. Entre as propostas dentro da FBC, destacam-se a expansão do conjunto de POIs, o balanceamento da cobertura entre categorias e o uso de pesos ajustáveis para os atributos dos itens. A médio prazo, sugere-se a adoção de LSA por meio de SVD, visando capturar relações semânticas mais profundas.

Já para futuras integrações de SR Híbrido, sugere-se o uso de LLMs, que possibilitará uma interpretação mais precisa das preferências, capturando nuances contextuais de forma mais eficaz. Isso aumentará significativamente a diversidade das recomendações e permitirá uma adaptação dinâmica às mudanças de interesse ao longo do tempo. Além de integração de abordagens colaborativas, como *Neighborhood-Based*, *Model-Based* e *Deep Auto Encoders*, para captar preferências implícitas ou comportamentais.

Referências

- Achakulvisut, T., Acuna, D. E., Ruangrong, T., and Kording, K. (2016). Science Concierge: A Fast Content-Based Recommendation System for Scientific Publications. *PLOS ONE*, 11(7):e0158423.
- Aggarwal, C. C. (2016). *Recommender Systems: The Textbook*. Springer, Cham, Switzerland.
- Anthis, J. R., Liu, R., Richardson, S. M., Kozłowski, A. C., Koch, B., Brynjolfsson, E., Evans, J., and Bernstein, M. S. (2025). Position: LLM Social Simulations Are a Promising Research Method. In *Proceedings of the 42nd International Conference on Machine Learning*, PMLR 267. PMLR. Vancouver, Canada.
- Borràs, J., Moreno, A., and Valls, A. (2014). Intelligent tourism recommender systems: a survey. *Expert Systems with Applications*, 41(16):7370–7389.
- Campos, J., de Roldão, I., Alves, P., Dias, M., Moura, B., Freire, K., dos Santos, N. C., Freitas, P., Braga, A., Galvão, B., Simões, F., Jarocki, I., Paulo, J., Almeida, L., Marques, P., Rayane, S., Xavier, S., Instituto de Assessoria para o Desenvolvimento Humano (IADH), and Rede Nacional de Experiências e Turismo Criativo (RECRIA) (2022). Plano de turismo criativo 2022-2024.
- Castells, P. and Jannach, D. (2023). Recommender systems: a primer. In Alonso, O. and Baeza-Yates, R., editors, *Advanced Topics for Information Retrieval*. ACM Press, New York, NY, USA.
- Costa, I. (2024). Carnaval do Recife movimentou R\$2,4 bilhões e recebeu mais de 3,4 milhões de foliões; prefeitura cogita juntar data magna aos dias oficiais de folia em 2025. Portal G1.
- Creswell, J. W. and Creswell, J. D. (2021). *Projeto de pesquisa: Métodos qualitativo, quantitativo e misto*. Penso, Porto Alegre, RS, Brasil, 5 edition.
- Deshpande, M. and Karypis, G. (2004). Item-based top-N recommendation algorithms. *ACM Transactions on Information Systems*, 22(1):143–177.
- Ferreira, D. I. F. and Ramos, C. (2024). Uma análise estratégica com base na matriz SWOT: a aplicabilidade do projeto TheRoute a Pernambuco. In *Anais do 15º Congresso de Administração, Sociedade e Inovação (CASI)*.
- Fonseca, D. (2020). Pernambuco lança passaporte para incentivar turismo no estado entre moradores e turistas. Portal G1.
- Gao, H., Tang, J., Hu, X., and Liu, H. (2015). Context-aware point of interest recommendations on location-based social networks. In *Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence*, pages 1721–1727, Austin, TX, USA. AAAI Press.
- Gao, M., Liu, K., and Wu, Z. (2010). Personalisation in web computing and informatics: the techniques, applications, and future research. *Information Systems Frontiers*, 12:607–629.
- Gomede, E., de Barros, R. M., and de Souza Mendes, L. (2021). Deep auto encoders to adaptive e-learning recommender system. *Computers and Education: Artificial Intelligence*, 2:100009.

- Jesse, M., Bauer, C., and Jannach, D. (2023). Intra-list similarity and human diversity perceptions of recommendations: The details matter. *User Modeling and User-Adapted Interaction*, 33:769–802.
- Kang, E., Kim, H., and Cho, J. (2006). Personalization method for tourist point of interest (poi) recommendation. In Gabrys, B., Howlett, R. J., and Jain, L. C., editors, *Knowledge-Based Intelligent Information and Engineering Systems*, volume 4251 of *Lecture Notes in Computer Science*, pages 392–400, Berlin, Germany. Springer.
- Li, Y., Liu, K., Wang, S., and Cambria, E. (2021). Recent developments in recommender systems: A survey. *Journal of \LaTeX Class Files*, 14(8).
- Magara, M. B., Ojo, S. O., and Zuva, T. (2018). A Comparative Analysis of Text Similarity Measures and Algorithms in Research Paper Recommender Systems. In *Proceedings of the 2018 Conference on Information Communications Technology and Society (ICTAS)*. IEEE. Pretoria, South Africa.
- Morais, I., Mendonça, E., and Santos, E. (2022). Novas formas de fazer turismo: desde a prática às políticas na construção do plano de turismo criativo do recife (pernambuco-brasil). *Interações*, 23(3):669–684.
- Kaiser, S. and Ali, R. (2018). Text mining: use of TF-IDF to examine the relevance of words to documents. *International Journal of Computer Applications*, 181(1):25–29.
- Ries, E. (2012). *A startup enxuta: Como empreendedores atuais utilizam a inovação contínua para criar empresas extremamente bem-sucedidas*. Leya, São Paulo, SP, Brasil.
- Salton, G. and Buckley, C. (1988). Term-weighting approaches in automatic text retrieval. *Information Processing & Management*, 24(5):513–523.
- Sarkar, J. P., Majumder, A., Panigrahi, C., Roy, S., and Pati, B. (2023). Tourism recommendation system: a survey and future research directions. *Multimedia Tools and Applications*, 82:8983–9027.
- Sondur, S. D. and Chigadani, A. P. (2016). Similarity measures for recommender systems: a comparative study. *Journal for Research*, 2(3):76–80.
- Trigo, L. G. G. (2020). Viagens e turismo: dos cenários imaginados às realidades disruptivas. *RBTUR*, 14(3):1–13.
- Valença, J. (2022). Passaporte pernambuco: saiba onde e como conseguir o caderno turístico.
- Wazlawick, R. S. (2009). *Metodologia de pesquisa para ciência da computação*. Elsevier, Rio de Janeiro, RJ, Brasil, 6 edition.
- Zhao, S., King, I., and Lyu, M. R. (2016). A survey of point-of-interest recommendation in location-based social networks. *ACM Computing Surveys*, 49(2):Artigo 25.
- Zhao, Z., Fan, W., Li, J., Liu, Y., Mei, X., Wang, Y., Wen, Z., Wang, F., Zhao, X., Tang, J., and Li, Q. (2024). Recommender systems in the era of large language models (LLMs). *IEEE Transactions on Knowledge and Data Engineering*, 36(11):6889–6907.